REMARKS

Reconsideration of the above referenced application in view of the following remarks is requested. In the specification, the abstract of the disclosure has been amended to correct a sentence fragment as required by the Office Action; the sentence in the second full paragraph on page 8 between line 17 and line 19 has been amended to correct an editorial error. In the claims, claims 1-3, 5, 10-11, 13, 16, 25, and 27 have been amended. Claims 4 and 6 have been cancelled. Existing claims 1-3, 5, 7-29 remain in the application.

ARGUMENT

Claim Rejections – 35 U.S.C. § 112

Claims 13 and 14 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 13 has been amended to depend upon claim 12. The limitation "the controller" in claim 13 can now find its antecedent basis in claim 12. Thus, the 35 U.S.C. § 112 rejection of claims 13-14 is now moot.

Claim Rejections – 35 U.S.C. § 102

Claims 1-6, 8-27, and 29 are rejected under 35 U.S.C. 102(b) as being anticipated by Roberts, et al. (U.S. Patent No. 5,027,406) (hereinafter Roberts).

The above-identified application discloses a speech recognition system whose operations can be automatically tuned based on feedback data. The system includes at

least one speech recognizer to produce output signals from audio input signals and a feedback module to collect feedback data generated from conversion of the audio input signals to output signals. The feedback data may be used to modify speech models and grammar files of the speech recognizer, either in real time or off-line, or both, so that the recognizer can be tuned to have improved performance for an application. The feedback data may be collected either explicitly or implicitly. The speech recognition system disclosed in the present application may be used in either command and control applications or dictation applications, or both.

Roberts discloses a method for creating word models for a large vocabulary, natural language dictation system. Using this method, a user with limited typing skills can create documents with little or no advance training of word models. As the user is dictating, the system displays a list of words in the active vocabulary which best match a spoken word. By keyword or voice command, the user may choose the correct word from the list or may choose to edit a similar word if the correct word is not on the list. When the correct word is chosen by the user, the speech sample is used to create or update an acoustic model for the word, without further intervention by the user. As the system is used, it also constantly updates its statistical language model. The system gets more and more word models and keeps improving its performance the more it is used.

Turning now to claim 1, this claim has been amended. The amended claim 1 recites a controller adaptable to modify the speech models and the grammar files based on the feedback data to improve the performance of the at least one recognizer, in addition to other limitations recited in the original claim 1. In rejecting original claims 2,

3, 12, 13, 15, 21, and 26, the Examiner asserted, "*Roberts et al.* discloses a block diagram of a computer program for coordinating output of text by speech recognition ('production of the output signals') and editing by selection command 125, edit menu choice commands 126, and letter commands 127 ('adaptable to provide the feedback data to the recognizer') (Figure 1); the computer program is 'a controller'." This is an erroneous assertion. Fig. 1 of Roberts is a schematic block diagram of the functional steps performed in a preferred embodiment of the present invention designed to recognize separately spoken words (see col. 5, lines 52-55 of Roberts). The computer program that performs functions shown in Fig. 1 includes speech recognition (either in EDITMODE (122) or TEXTMODE (123)), and many other functions that support recognition function such as application-program interfacing (e.g., 274, 802, etc.), storing (e.g., 208, 216), and loading (e.g., 100). The program implements an integral speech recognition system. In marked contrast, the controller recited in claim 1 does not perform recognition functions because those functions are specifically performed by the recognizer, which is another component recited in claim 1. Thus, the computer program illustrated in Fig. 1 of Roberts is not a controller as recited in claim 1. Roberts does not teach or suggest the controller as recited in claim 1.

Additionally, Roberts only discloses a dictation system that does not use grammar files. In marked contrast, amended claim 1 explicitly recites grammar file elements. In rejecting original claim 4, the Examiner made an erroneous assertion that the language model used in Roberts is a grammar file. As the Examiner pointed out in the Office Action, the language model includes probabilities of a word followed by another word. Such a language model is part of speech models disclosed in the

present application. A grammar file, however, is a file different from a language model. It is typically used by a speech recognizer for a command and control application to specify a variety of speech patterns recognizable by the recognizer. See page 3, lines 16 to 20 of the specification of the present application. A speech recognizer for a dictation application typically does not use grammar files partly because there are too many possible speech patterns to be included in a grammar file for a dictation application. An example of a grammar file is shown on page 10, lines 8 to 14. Obviously, such a grammar file is different from a language model which contains probabilities of word sequences (e.g., word 2 given word 1) obtained based on statistics of a large speech training corpus before a recognizer is put into use (see col. 13, lines 44 to 60 of Roberts). The language model disclosed in Roberts is totally different from the grammar file recited in claim 1. Nowhere does Roberts teach or suggest grammar files as recited in claim 1.

Because Roberts does not teach or suggest the controller or grammar files recited in amended claim 1, this independent claim is not anticipated by Roberts and is thus patentable over Roberts.

As to independent claim 11, it has been amended to recite the grammar files element. Based on the forgoing reasons presented to support the patentability of amended claim 1, Roberts does not teach or suggest grammar files element recited in amended claim 11. Thus, amended claim 11 is patentable over Roberts.

Regarding independent claims 16 and 25, they have been amended to correct editorial errors. In rejecting original claims 16 and 25, the Examiner cited column 8, line 58 to column 9, line 7 of Roberts as teaching the claimed limitation: estimating a

correctness measure wherein the correctness measure expresses if the output signal is a correct representation of the audio input signal. The cited portion of Roberts discloses how to compare the sequence of individual frames of an unknown input spoken word with each of a plurality of acoustic word models and compute a score for each model indicating how the model matches the frame sequence of the input word. The words whose models generate the best scores are selected as the best candidates for the unknown input spoken word. A list of candidate words is shown for a user to choose.

In marked contrast, the correctness measure recited in claim 16 and 25 is not obtained by measuring an unknown input speech utterance with acoustic models. The claimed correctness measure is determined by a grammar file or by an explicit indication by an application or a user. The claimed correctness measure is not directly computed by comparing an aligned frame sequence of an unknown input utterance with acoustic word models. Normally, the correctness measure does not result in a list of candidates for a user to pick. Therefore, estimating a correctness measure in the present application is different from computing a score reflecting how close an input speech matches a word acoustic model as taught in Roberts. Roberts does not teach or suggest this claimed limitation.

Because Roberts does not teach or suggest the limitation of estimating a correctness measure recited in independent claims 16 and 25, these claims are patentable over Roberts.

Because independent claims 1, 11, 16, and 25 are patentable over Roberts, all claims dependent therefrom are also patentable over Roberts (e.g., existing claims 2-3, 5, 7-10, 12-15, 17-24, and 24-29).

Claim Rejections – 35 U.S.C. § 103

Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Roberts in view of Thelen, et al. (U.S. Patent No. 6,526,380) (hereinafter Thelen).

Thelen discloses a speech recognition system which includes a plurality of large vocabulary speech recognizers each being associated with a respective, different large vocabulary recognition model. Each of the recognition models is targeted to a specific part of a huge vocabulary. The system comprises a controller operative to direct the input pattern to a plurality of the speech recognizers and to select a recognized word sequence from the word sequences recognized by the plurality of speech recognizers. According to the portion of Thelen cited by the Examiner (col. 7, line 30 to col. 8, line 5) as well as the entire Thelen reference, it is clear that each speech recognizer can only recognize a portion of the vocabulary, in other words, the task of recognizing input utterances which may contain a huge vocabulary is distributed to multiple speech recognizers, each specializing in recognizing only a portion of the huge vocabulary.

In marked contrast, claim 7 recites multiple recognizers and a predictor to select a best performing recognizer from the multiple recognizers based upon the feedback data. The claimed multiple recognizers are not used to deal with a huge vocabulary specifically and each recognizer here is not specialized in only a portion of a vocabulary. Each recognizer may recognize any input utterance (i.e., any portion of the vocabulary) but probably with different performance. The claimed predictor helps to select one recognizer from the multiple recognizers based on the feedback data. Thus,

Thelen does not teach or suggest the claimed multiple recognizers and the claimed predictor.

Additionally, based on the arguments presented in traversing the 35 U.S.C. § 102 rejections, Roberts does not teach or suggest all of the limitations recited in amended independent claim 1, e.g., the controller element and the grammar files element. Claim 7 depends upon amended claim 1 and includes all the limitations recited in claim 1. Nowhere does Thelen teach or suggest those elements which are not taught or suggested by Roberts (e.g., the controller element and the grammar files element) either.

Because the combination of Roberts and Thelen does not teach or suggest all the limitations recited in claim 7, this claim is patentable over Roberts in view of Thelen.

Claim 28 is rejected under 35 U.S.C. § 103(a) as being unpatentable over Roberts in view of Ortega (U.S. Patent No. 6,157,910) (hereinafter Ortega).

Ortega discloses a method for updating an original speaker's speech file in a speech application. According to the method, when a user makes corrections to a document, the corrections for the document are logged into a file, which contains a log of corrections. The file log identifies changes to the language model and any new words added through correction. The user can take this log to another system. By invoking an Update Speech File command, for example from a menu, the system then reads the log and performs the same corrections to the original speaker's speech files as was performed on the other system.

According to the cited portion of Ortega cited by the Examiner (col. 1, line 44 to col. 2, line 6) as well as the entire Ortega reference, it is clear that Ortega discloses a method for how to apply a user's changes from one system to another system through a log file. The log file only contains corrections a user made to a document. Nowhere does Ortega disclose that the log file stores audio input signals. In marked contrast, the claimed limitation in claim 28 includes storing only those audio input signals for which the correction status indicates that a correction to the output signal was necessary. Additionally, the audio input signals are stored in the present application for the use of the current speech recognition system, not for another recognition system as disclosed in Ortega. Thus, Ortega does not teach or suggest the limitation recited in claim 28.

Additionally, based on the arguments presented in traversing the 35 U.S.C. § 102 rejections, Roberts does not teach or suggest all the limitations recited in amended independent claim 25, e.g., the element of estimating a correctness measure. Claim 28 depends upon amended claim 25 and includes all the limitations recited in claim 25. Nowhere does Ortega teach or suggest those elements which are not taught or suggested by Roberts (e.g., the element of estimating a correctness measure) either.

Because the combination of Roberts and Ortega does not teach or suggest all the limitations recited in claim 28, this claim is patentable over Roberts in view of Ortega.

CONCLUSION

In view of the foregoing, claims 1-3, 5, 7-29 are all in condition for allowance. If the Examiner has any questions, the Examiner is invited to contact the undersigned at (503) 264-8074. Early issuance of a Notice of Allowance is respectfully requested.

Respectfully submitted,

Date: 6/30/05

Steven P. Skabrat
Senior Attorney
Intel Corporation
Registration No. 36,279
(503) 264-8074

c/o Blakely, Sokoloff, Taylor & Zafman, LLP
12400 Wilshire Boulevard
Seventh Floor
Los Angeles, CA 90025-1026